



# Deep learning for the segmentation of preserved photoreceptors on *en face* optical coherence tomography in two inherited retinal diseases

ACNER CAMINO,<sup>1,3</sup> ZHUO WANG,<sup>2,3</sup> JIE WANG,<sup>1</sup> MARK E. PENNESI,<sup>1</sup> PAUL YANG,<sup>1</sup> DAVID HUANG,<sup>1</sup> DENGWANG LI,<sup>2</sup> AND YALI JIA<sup>1,\*</sup>

<sup>1</sup>Casey Eye Institute, Oregon Health and Science University, Portland, OR, 27239, USA

<sup>2</sup>Shandong Province Key Laboratory of Medical Physics and Image Processing Technology, School of Physics and Electronics, Shandong Normal University, Jinan, 250014, China

<sup>3</sup>These authors contributed equally to this manuscript

\*[jiaya@ohsu.edu](mailto:jiaya@ohsu.edu)

**Abstract:** The objective quantification of photoreceptor loss in inherited retinal degenerations (IRD) is essential for measuring disease progression, and is now especially important with the growing number of clinical trials. Optical coherence tomography (OCT) is a non-invasive imaging technology widely used to recognize and quantify such anomalies. Here, we implement a versatile method based on a convolutional neural network to segment the regions of preserved photoreceptors in two different IRDs (choroideremia and retinitis pigmentosa) from OCT images. An excellent segmentation accuracy (~90%) was achieved for both IRDs. Due to the flexibility of this technique, it has potential to be extended to additional IRDs in the future.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

**OCIS codes:** (170.4500) Optical coherence tomography; (170.4470) Ophthalmology; (100.6890) Three-dimensional image processing; (170.1610) Clinical applications.

## References and links

1. P. J. Francis, "Genetics of inherited retinal disease," *J. R. Soc. Med.* **99**(4), 189–191 (2006).
2. G. Liu, X. Liu, H. Li, Q. Du, and F. Wang, "Optical coherence tomographic analysis of retina in retinitis pigmentosa patients," *Ophthalmic Res.* **56**(3), 111–122 (2016).
3. R. Sanchez-Alcudia, M. Garcia-Hoyos, M. A. Lopez-Martinez, N. Sanchez-Bolivar, O. Zurita, A. Gimenez, C. Villaverde, L. Rodrigues-Jacy da Silva, M. Corton, R. Perez-Carro, S. Torriano, V. Kalatzis, C. Rivolta, A. Avila-Fernandez, I. Lorda, M. J. Trujillo-Tiebas, B. Garcia-Sandoval, M. I. Lopez-Molina, F. Blanco-Kelly, R. Riveiro-Alvarez, and C. Ayuso, "A comprehensive analysis of choroideremia: from genetic characterization to clinical practice," *PLoS One* **11**(4), e0151943 (2016).
4. T. Rosenberg, M. Haim, A.-M. Hauch, and A. Parving, "The prevalence of Usher syndrome and other retinal dystrophy-hearing impairment associations," *Clin. Genet.* **51**(5), 314–321 (1997).
5. M. A. Genead, G. A. Fishman, E. M. Stone, and R. Allikmets, "The natural history of Stargardt disease with specific sequence mutation in the ABCA4 Gene," *Invest. Ophthalmol. Vis. Sci.* **50**(12), 5867–5871 (2009).
6. R. Sitorus, M. Preising, and B. Lorenz, "Causes of blindness at the 'Wiyata Guna' School for the Blind, Indonesia," *Br. J. Ophthalmol.* **87**(9), 1065–1068 (2003).
7. P. Goodwin, "Hereditary retinal disease," *Curr. Opin. Ophthalmol.* **19**(3), 255–262 (2008).
8. R. Syed, S. M. Sundquist, K. Ratnam, S. Zayit-Soudry, Y. Zhang, J. B. Crawford, I. M. MacDonald, P. Godara, J. Rha, J. Carroll, A. Roorda, K. E. Stepien, and J. L. Duncan, "High-resolution images of retinal structure in patients with choroideremia," *Invest. Ophthalmol. Vis. Sci.* **54**(2), 950–961 (2013).
9. M. Yung, M. A. Klufas, and D. Sarraf, "Clinical applications of fundus autofluorescence in retinal disease," *Int J Retina Vitreous* **2**(1), 12 (2016).
10. A. Oishi, K. Ogino, Y. Makiyama, S. Nakagawa, M. Kurimoto, and N. Yoshimura, "Wide-field fundus autofluorescence imaging of retinitis pigmentosa," *Ophthalmology* **120**(9), 1827–1834 (2013).
11. E. Garcia-Martin, I. Pinilla, E. Sancho, C. Almarcegui, I. Dolz, D. Rodriguez-Mena, I. Fuertes, and N. Cuenca, "Optical coherence tomography in retinitis pigmentosa: reproducibility and capacity to detect macular and retinal nerve fiber layer thickness alterations," *Retina* **32**(8), 1581–1591 (2012).

12. G. Staurengi, S. Sadda, U. Chakravarthy, and R. F. Spaide, "Proposed lexicon for anatomic landmarks in normal posterior segment spectral-domain optical coherence tomography," *Ophthalmology* **121**(8), 1572–1578 (2014).
13. D. G. Birch, Y. Wen, K. Locke, and D. C. Hood, "Rod sensitivity, cone sensitivity, and photoreceptor layer thickness in retinal degenerative diseases," *Invest. Ophthalmol. Vis. Sci.* **52**(10), 7141–7147 (2011).
14. G. Liu, H. Li, X. Liu, D. Xu, and F. Wang, "Structural analysis of retinal photoreceptor ellipsoid zone and postreceptor retinal layer associated with visual acuity in patients with retinitis pigmentosa by ganglion cell analysis combined with OCT imaging," *Medicine (Baltimore)* **95**(52), e5785 (2016).
15. D. Mukherjee, E. M. Lad, R. R. Vann, S. J. Jaffe, T. E. Clemons, M. Friedlander, E. Y. Chew, G. J. Jaffe, and S. Farsiu; MacTel Study Group, "Correlation between macular integrity assessment and optical coherence tomography imaging of ellipsoid zone in macular telangiectasia type 2," *Invest. Ophthalmol. Vis. Sci.* **58**(6), BIO291 (2017).
16. W. Zhu, H. Chen, H. Zhao, B. Tian, L. Wang, F. Shi, D. Xiang, X. Luo, E. Gao, L. Zhang, Y. Yin, and X. Chen, "Automatic three-dimensional detection of photoreceptor ellipsoid zone disruption caused by trauma in the OCT," *Sci. Rep.* **6**(1), 25433 (2016).
17. Z. Wang, A. Camino, M. Zhang, J. Wang, T. S. Hwang, D. J. Wilson, D. Huang, D. Li, and Y. Jia, "Automated detection of photoreceptor disruption in mild diabetic retinopathy on volumetric optical coherence tomography," *Biomed. Opt. Express* **8**(12), 5384–5398 (2017).
18. Z. Wang, A. Camino, A. M. Hagag, J. Wang, R. G. Weleber, P. Yang, M. E. Pennesi, D. Huang, D. Li, and Y. Jia, "Automated detection of preserved photoreceptor layer on optical coherence tomography in choroideremia based on machine learning," *J. Biophotonics*, doi:10.1002/jbio.201700313(2018).
19. K. Xue, M. Oldani, J. K. Jolly, T. L. Edwards, M. Groppe, S. M. Downes, and R. E. MacLaren, "Correlation of optical coherence tomography and autofluorescence in the outer retina and choroid of patients with choroideremia," *Invest. Ophthalmol. Vis. Sci.* **57**(8), 3674–3684 (2016).
20. N. Jain, Y. Jia, S. S. Gao, X. Zhang, R. G. Weleber, D. Huang, and M. E. Pennesi, "Optical coherence tomography angiography in choroideremia: correlating choriocapillaris loss with overlying degeneration," *JAMA Ophthalmol.* **134**(6), 697–702 (2016).
21. S. S. Gao, G. Liu, D. Huang, and Y. Jia, "Optimization of the split-spectrum amplitude-decorrelation angiography algorithm on a spectral optical coherence tomography system," *Opt. Lett.* **40**(10), 2305–2308 (2015).
22. Y. Jia, O. Tan, J. Tokayer, B. Potsaid, Y. Wang, J. J. Liu, M. F. Kraus, H. Subhash, J. G. Fujimoto, J. Hornegger, and D. Huang, "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography," *Opt. Express* **20**(4), 4710–4725 (2012).
23. M. F. Kraus, B. Potsaid, M. A. Mayer, R. Bock, B. Baumann, J. J. Liu, J. Hornegger, and J. G. Fujimoto, "Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns," *Biomed. Opt. Express* **3**(6), 1182–1199 (2012).
24. R. Zhao, A. Camino, J. Wang, A. M. Hagag, Y. Lu, S. T. Bailey, C. J. Flaxel, T. S. Hwang, D. Huang, D. Li, and Y. Jia, "Automated drusen detection in dry age-related macular degeneration by multiple-depth, *en face* optical coherence tomography," *Biomed. Opt. Express* **8**(11), 5049–5064 (2017).
25. M. Zhang, J. Wang, A. D. Pechauer, T. S. Hwang, S. S. Gao, L. Liu, L. Liu, S. T. Bailey, D. J. Wilson, D. Huang, and Y. Jia, "Advanced image processing for optical coherence tomographic angiography of macular diseases," *Biomed. Opt. Express* **6**(12), 4661–4675 (2015).
26. R. F. Spaide and C. A. Curcio, "Anatomical correlates to the bands seen in the outer retina by optical coherence tomography: literature review and model," *Retina* **31**(8), 1609–1619 (2011).
27. A. Camino, Y. Jia, G. Liu, J. Wang, and D. Huang, "Regression-based algorithm for bulk motion subtraction in optical coherence tomography angiography," *Biomed. Opt. Express* **8**(6), 3053–3066 (2017).
28. A. Vedaldi and K. Lenc, "MatConvNet: Convolutional Neural Networks for MATLAB," in *Proceedings of the 23rd ACM international conference on Multimedia*, (ACM, Brisbane, Australia, 2015), pp. 689–692.
29. N. Jain, Y. Jia, S. S. Gao, X. Zhang, R. G. Weleber, D. Huang, and M. E. Pennesi, "Optical coherence tomography angiography in choroideremia: correlating choriocapillaris loss with overlying degeneration," *JAMA Ophthalmol.* **134**(6), 697–702 (2016).
30. L. Fang, D. Cuneffare, C. Wang, R. H. Guymier, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express* **8**(5), 2732–2744 (2017).

## 1. Introduction

Inherited retinal degenerations (IRDs) are caused by mutations in genes important for retinal function and cause progressive retinal degeneration. The most common IRD is retinitis pigmentosa (prevalence of 1 in 3,000–4,000 people) [1, 2], but other common IRDs include: choroideremia (estimated 1 in 50,000) [3], Usher syndrome (approximately 1 in 20,000) [4], Stargardt disease (1 in 8,000–10,000) [5], Leber amaurosis (2–3 in 100,000) [6] and others [7]. Since IRDs progressively lead to blindness, it is of great importance to monitor the integrity of photoreceptors in routine follow up visits and during gene therapy.

Currently, several imaging technologies encompassing fundus photography [8], fundus autofluorescence [9, 10] and optical coherence tomography (OCT) [2, 8, 11] are used for assessment of disease progression in the clinical practice. Since OCT is the only one to provide depth-resolved information of retinal tissue, it is the most solid existing technology for imaging and quantification of photoreceptor preservation.

In OCT images, the second hyper-reflective layer of the outer retina, identified as the ellipsoid zone (EZ) of the photoreceptors [12], is the structure most suitable to assess photoreceptor damage [13]. Numerous image processing techniques have been reported in recent years to detect and quantify the extent of EZ damage in IRDs [14], macular telangiectasia [15] and ocular trauma [16]. Previously, we have developed *en face* methods that use OCT images to detect EZ loss in mild diabetic retinopathy by fuzzy logic [17] and choroideremia by a random forest classifier [18]. However, the pattern of photoreceptor integrity can present differently in each retinal pathology. For example, with retinitis pigmentosa the best strategy is to detect the preserved EZ boundary since the degeneration starts in the mid-periphery and constricts centrally to leave a round-shaped “island” of preserved EZ centered at the fovea [2]. For other diseases, such as Stargardt Dystrophy, where photoreceptor atrophy starts centrally, it is more feasible to detect EZ loss. The pattern of EZ atrophy can present with complex shapes such as with choroideremia, which shows initial loss in the periphery of the macula, scalloped edges [19] and outer retinal tubulations [20]. Consequently, image processing methods developed targeting a certain disease assuming certain ad hoc rules are not generalizable and typically do not perform as well for patients with a different IRD.

With the purpose of developing a single method that is adaptable to different retinal conditions, we have implemented a deep learning platform that can be trained for more than one IRD (herein, retinitis pigmentosa and choroideremia) to detect the areas of preserved EZ. Our approach uses a segmentation method consisting of sliding-window binary classification of OCT B-scan sections by a convolutional neural network (CNN). In the context of deep learning, the segmentation problem is that of finding the pixels that belong to a certain semantic class that the network has been trained to recognize (e.g. defect tissue vs. healthy tissue). Here, we use a CNN trained from B-scan patches enclosing sections of the EZ, each of which is labeled based on the appearance of *en face* images at the patch’s central A-line position. Further bimodal thresholding of probability maps by an Otsu scheme and morphological operations provided binary maps of the segmented preserved photoreceptor areas with high accuracy compared to manual segmentation by an expert grader.

## 2. Materials and methods

### 2.1 Study population

Twenty subjects diagnosed with choroideremia and twenty-two diagnosed with retinitis pigmentosa were recruited from the Ophthalmic Genetics clinic at the Casey Eye Institute at the Oregon Health & Science University (OHSU). The protocol was approved by the Institutional Review Board/Ethics Committee of OHSU and the research adhered to the tenants of the Declaration of Helsinki.

### 2.2 Data acquisition

Macular scans covering a 6 mm × 6 mm area were acquired by a 70-kHz, 840-nm-wavelength spectral-domain OCT system (Avanti RTVue-XR, Optovue Inc.) within 2.9 seconds. The AngioVue version 2016.2.0.35 software was used to acquire optical coherence tomography angiography (OCTA) scans. In the fast transverse scanning direction, 304 A-scans were sampled to form a B-scan and two repeated B-scans were acquired at each lateral location. A total of 304 locations were scanned in the slow transverse direction to form a 3D data cube. Axial resolution in AngioVue is 5 μm but digital pixel sampling is 3 μm. Structural OCT data

was obtained by averaging the two repeated B-scans at each location and OCTA data was generated by the split-spectrum amplitude-decorrelation (SSADA) algorithm [21, 22]. In order to remove microsaccadic artifacts and improve the signal-to-noise ratio of images, two sets of volumetric data were acquired at orthogonal scanning directions, registered and merged by motion correction technology (MCT<sup>TM</sup>) [23].

### 2.3 Image processing method summary

The algorithm is divided into five parts: pre-processing, manual grading, patch extraction, training of a neural network using patches and post-processing [Fig. 1]. The pre-processing step uses the segmentation of the Bruch's membrane interface to generate a flattened B-scan. This interface was chosen because it is preserved in both diseases and can be reliably segmented [24]. Confounding shadows projected onto the EZ by the large vessels on inner retina were removed in this step. Then, squared patches containing the EZ were extracted from B-scans and used to train a CNN. Finally, median filtering, an Otsu thresholding scheme and morphological processing were used to extract the two-dimensional image of preserved EZ in the post-processing step for the test data set. The following sections will describe these processes in detail. The algorithm was implemented with custom software written in Matlab 2017a (Mathworks, Natick, MA) and the MatConvNet platform (<http://www.vlfeat.org/matconvnet/>).

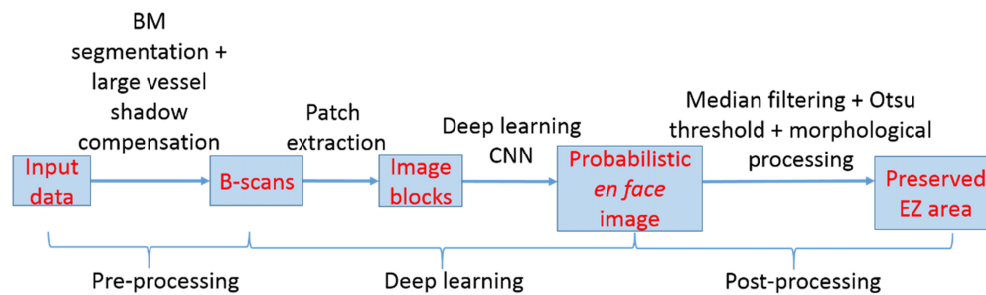


Fig. 1. Flow chart of the preserved ellipsoid zone detection algorithm. Patches from B-scans containing the ellipsoid zone region are used to train convolutional neural network (CNN) for classification of each patch's central A-line position as preserved or loss of photoreceptors. BM – Bruch's membrane

### 2.4 Pre-processing

The inner limiting membrane (ILM), outer boundary of the outer plexiform layer (OPL) and Bruch's-membrane/choriocapillaris interfaces were segmented by a method based on directional graph search [25]. Rectangular sections of the B-scans were defined between the Bruch's membrane and 33 pixels above it. The region enclosed within this flattened B-scan section includes the EZ, a section of the layers above it (Myoid zone, external limiting membrane and outer nuclear layer) [12, 26] and the hyper-reflective layers below it (interdigitation zone and retinal pigment epithelium).

Since the EZ is a hyper-reflective layer on OCT images, photoreceptor loss areas can be identified by their lower reflectance. However, confounding shadow artifacts caused by the absorption of optical signal at the superficial large vessels need to be recognized and removed [Fig. 2(A-B), red arrows]. A mask of the inner retinal large vessels was constructed by a method reported previously [27]. Briefly, the *en face* angiogram of the inner retinal flow is generated by maximum projection of decorrelation values between ILM and OPL. Then, amplitude thresholding is applied, followed by morphological opening and a Gaussian convolution filter. The A-lines located at the positions recognized by the large vessel mask hold unreliable EZ layer information and were corrected by retrieving the information contained in its neighborhood. Specifically, for each C-scan contained in the rectangular



section of a B-scan [Fig. 2(B)], the reflectance values of the A-lines affected by shadows were substituted by the mean reflectance value of the pixels not contained in the large vessel mask within a circle of 10-pixel radius [Fig. 2(C)].

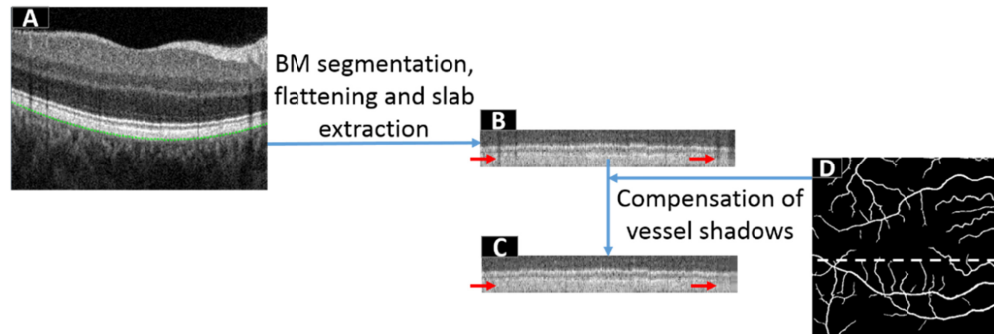


Fig. 2. Correction of shadow artifacts cast by large vessels onto outer retinal layers. A representative B-scan (A) is segmented at the Bruch's membrane (BM) and a segment between the BM and the C-scan located 33 pixels above the BM is generated (B). Shadow artifacts (red arrows) are corrected in (C) by averaging values in the vicinity within a circle of 10-pixel radius not contained in the large vessel mask (D). The dashed white line in (D) indicates the position of the representative B-scan.

## 2.5 Manual classification

In order to train a neural network to recognize preserved EZ from EZ loss, we divided the available data into train and test data sets, manually classified patches generated in Fig. 2 and assigned the corresponding label to the data used in the training stage. It is very hard for a human grader to classify patches of B-scans with confidence. Rather, manual segmentation was performed on *en face* images, which are more easily interpretable. In order to train for retinitis pigmentosa, a thickness map of the area between EZ and Bruch's membrane is a good feature for human graders to differentiate healthy from diseased areas with confidence. The inner boundary of EZ layer was segmented and the thickness map showed larger values at the positions where the EZ is preserved [Fig. 3]. Then an experienced, masked grader segmented the preserved EZ area using the thickness maps [Fig. 3]. It is challenging to provide reliable automatic segmentation of the EZ layer in choroideremia. Therefore, the mean projection of the slab generated between 8 to 16 pixels above the Bruch's membrane interface was used to approximate the location of the EZ layer, as performed previously [17] and assess photoreceptor integrity [Fig. 4]. In choroideremia, the healthy EZ area can be either partially or completely preserved. Partially preserved EZ has suffered a certain degree of damage but still contain functioning photoreceptors and hence, is more hyper-reflective than the EZ loss region but not as bright as the completely preserved EZ. As reported previously by us, the partially preserved EZ surrounds the completely preserved EZ and has a sharp contrast boundary with the EZ loss area, which is apparent to the manual grader [Fig. 4]. No distinction was made between partially and completely preserved EZ for grading purposes, and both were assigned to the same label.

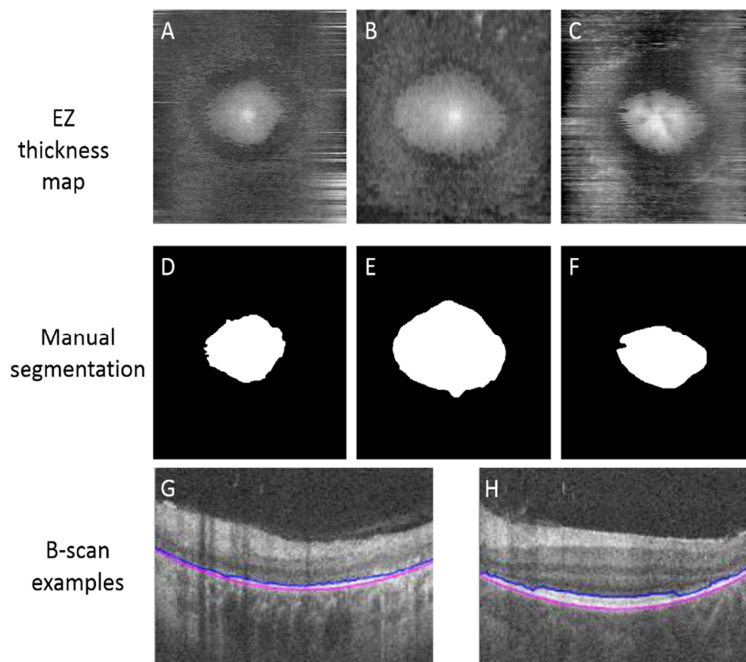


Fig. 3. The thickness maps between EZ and Bruch membrane of three retinitis pigmentosa subjects (A-C) and their corresponding manual grading (D-F). EZ – ellipsoid zone. G represents a B-scan example with total EZ loss. H represents a B-scan example with only peripheral EZ loss. Pink line is the Bruch's membrane. Blue line is the upper EZ boundary.

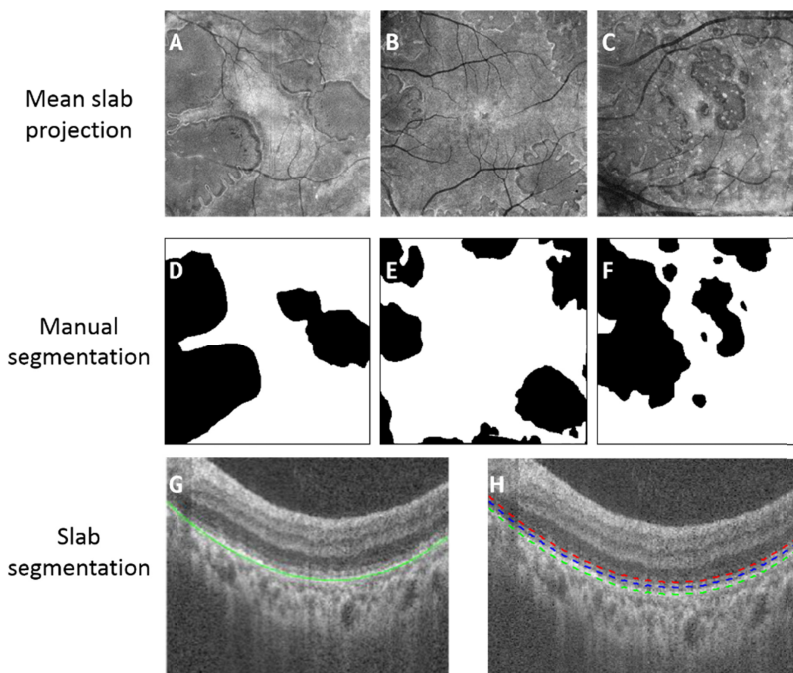


Fig. 4. The *en face* mean projection of the reflectance values within the 8-pixel-width slab for three choroideremia subjects (A-C) and their corresponding manual grading (D-F). Green line in representative B-scan in G and H represent the Bruch's membrane segmentation. Dashed lines in H represent the slab defined for *en face* projection.

## 2.6 Patch extraction

Squared patches of  $33 \times 33$  pixel size [Fig. 5] were extracted from the B-scan segments generated in Fig. 2(C). Each patch was then labeled as either EZ loss or preserved EZ conforming to the label assigned during manual grading to the  $(x, y)$  position of its central A-line, and fed to the CNN in the training stage. For the first and last 16 A-lines in each B-scan, their corresponding  $33 \times 33$ -pixels patch was completed by padding the missing lines with the B-scan's first or last A-line accordingly. For retinitis pigmentosa, the ratio of preserved EZ patches to EZ loss patches was 1.6:1 whereas for choroideremia the ratio was 2.1:1.

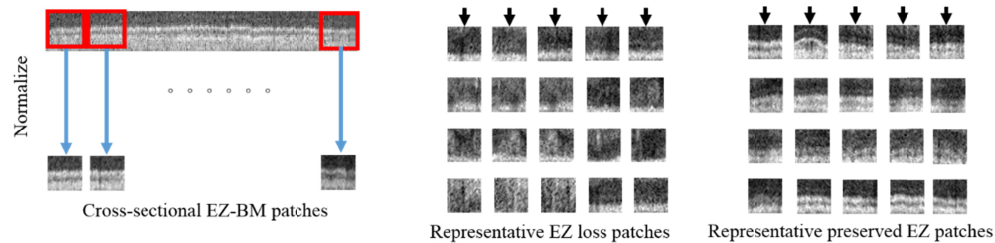


Fig. 5. Patch extraction process. Patches of size  $33 \times 33$  pixels are extracted from segments generated in Fig. 2(C) in order to train the CNN. Each patch is classified as ellipsoid zone (EZ) loss or preserved EZ, according to the manual grading assigned to the position  $(x, y)$  of its central A-line, indicated by black arrows. Preserved EZ patches show two distinct hyper-reflective boundaries: EZ and retinal pigment epithelium (RPE). Partially preserved and completely preserved EZ were assigned the same label on training stage.

## 2.7 En face preserved EZ segmentation by deep learning

The generic architecture of a CNN classifier is a sequential repetition of convolutional, activation and pooling layers, followed by a fully connected layer and soft-max classification. The convolutional layers consist of filter banks that use convolution of different spatial kernels to extract certain image features. The neurons in the convolutional layer have local receptive fields (i.e. they are not fully connected to the preceding layer, but to a subset of its neurons), convolve across the preceding layer with a pre-defined stride and share the same weight across the layer. After convolution, a three dimensional volume of feature maps is generated, containing the same number of feature maps as filters used. Pooling layers then compress the size of the feature maps, simplifying the network computation complexity. In addition to these two types of layers, activation by ReLU or Sigmoid function is used to introduce the nonlinearity to the neural network model. After these layers are stacked several times, a fully connected layer follows and the soft-max operation typically performs the task of classification.

Here, we train the CNN architecture in Fig. 6 pre-defined in the MatConvNet platform [28] to classify B-scan patches and assign its labels to the  $(x, y)$  position of the patch's central A-line. This network's architecture contains three convolutional layers, three pooling layers (one max pooling and two average pooling), ReLU activation, two fully connected layers and soft-max classification. The CNN was initialized by its default hyper-parameters and re-trained on B-scan patches to adapt its weights and biases to the specific classification task at hand. The network was trained over 45 epochs, with early stopping condition if the RMSE of validation set worsens for five successive epochs. Batch size was 100. L2 regularization with a weight decay of 0.0001 was used. Base learning rate was 0.05 for the first 30 epochs, reduced to 0.005 for the next 10 epochs and to 0.0005 for the last 5 epochs. The output was a  $1 \times 2$  vector containing the probabilities of patch belonging to either category (preserved EZ or EZ loss). The network was run on a desktop computer with an Intel CPU, 16 GB of RAM and a NVIDIA GPU (GeForce, Quadro K420) with 1GB VRAM.

Ten eyes with retinitis pigmentosa and ten eyes with choroideremia were used to train the network for each IRD separately. For each eye a volumetric scan with 304 B-scans was

available, from which one out of every 30 B-scans was chosen for training, for a total of 300 B-scans per IRD. From the 81600 image patches contained in the training set of B-scans, 61200 were selected for the training group (75%), 20400 for the validation group and performance was evaluated by 4-fold cross-validation. Image size at the input layer was  $33(x) \times 33(z) \times 1(y)$  pixels. Choroideremia and retinitis pigmentosa databases were trained separately by the same architecture.

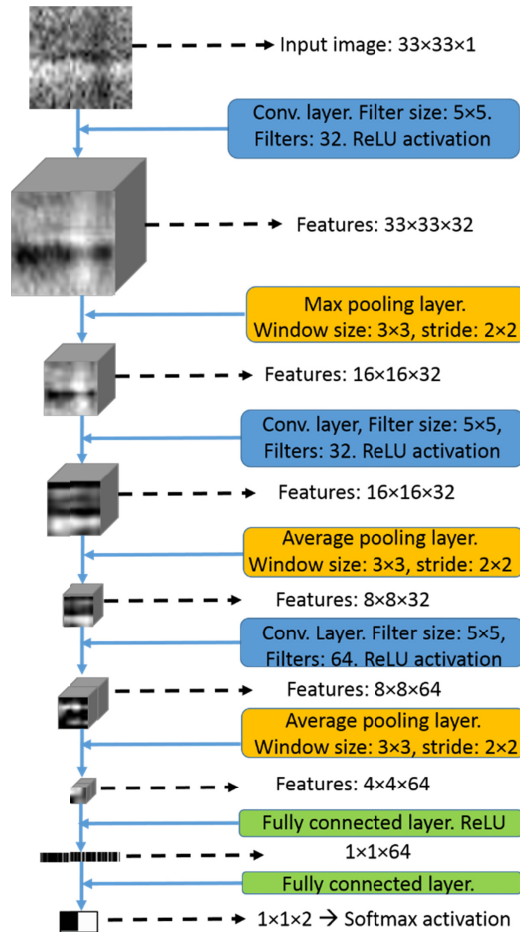


Fig. 6. Illustration of the convolutional neural network patch classification framework. Three convolutional layers with ReLU activation, three pooling layers and two fully connected layers are stacked together and trained with labeled patches of B-scan segments from the choroideremia and retinitis pigmentosa data sets separately.

## 2.8 Post-processing

A total of 92416 ( $304 \times 304$ ) patches were extracted from each scan under scrutiny and fed into the CNN model for A-line-wise classification, generating a probabilistic *en face* image of the preserved EZ region. Then, a median filter of kernel  $[10 \times 10]$  pixels was applied to remove noise. Otsu thresholding was used to identify a threshold from the bimodal histogram distribution of the probabilistic image. Morphological processing was applied on the resulting binary image to remove isolated areas smaller than 100 connected pixels.

### 3. Results

The CNN was applied on a set of 10 eyes with choroideremia, 12 eyes with retinitis pigmentosa and 5 healthy eyes, not including the eyes used in the training stage. The software completed the segmentation of any scan in the test data set within an average of 67 seconds.

The Jaccard similarity index, defined as the intersection divided by the union of the set of automatically segmented pixels  $S_R$  and manually segmented pixels  $S_M$

$$J(S_R, S_M) = \frac{|S_R \cap S_M|}{|S_R \cup S_M|} \quad (1)$$

was used to assess the accuracy of the segmentation. Its values range between 0 and 1, with large values meaning good similarity. Good qualitative correspondence with manual grading was observed for both diseases [Figs. 7-8]. Small isolated areas detected in the probability maps due to segmentation inaccuracies [Fig. 7, subjects 2 and 4; Fig. 8, subjects 3 and 4) were properly removed by the post-processing step. The similarity of automatic segmentation to manual grading was  $0.894 \pm 0.102$  (mean  $\pm$  pooled standard deviation) for retinitis pigmentosa and  $0.912 \pm 0.055$  for choroideremia.

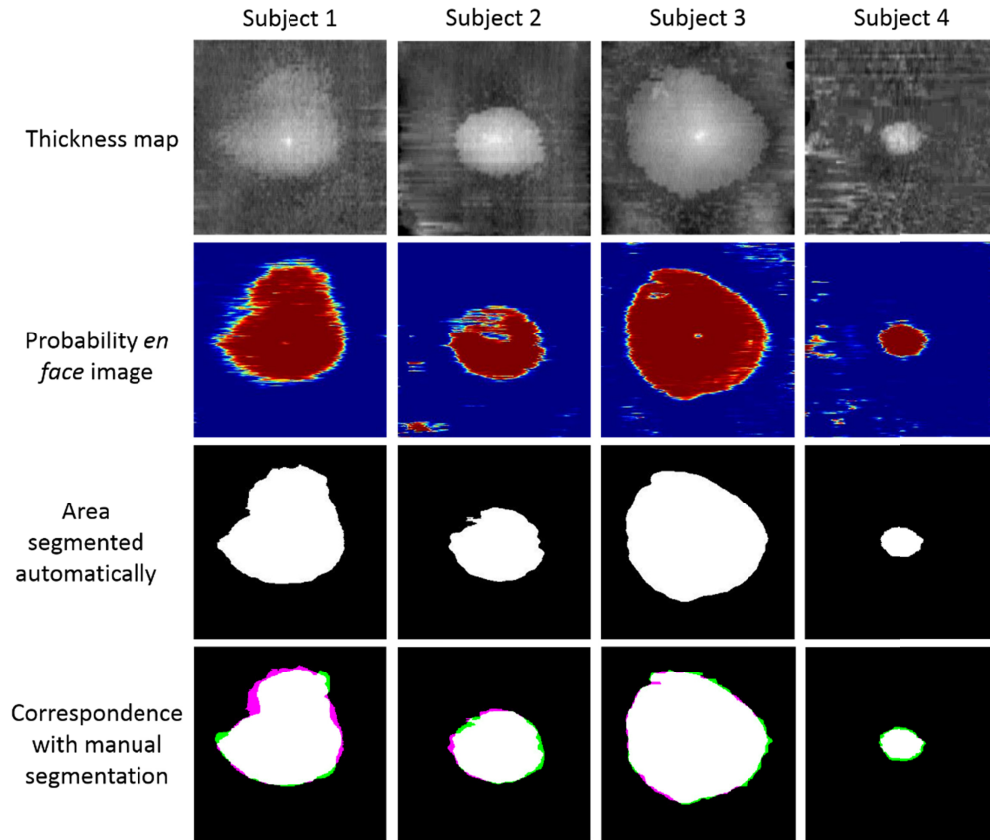


Fig. 7. Preserved ellipsoid zone areas detected automatically by the deep learning method and their correspondence with manual segmentation for 4 subjects with retinitis pigmentosa. Areas in white show agreement between manual segmentation and the CNN classifier, whereas areas in green are false negatives and areas in purple are false positives.



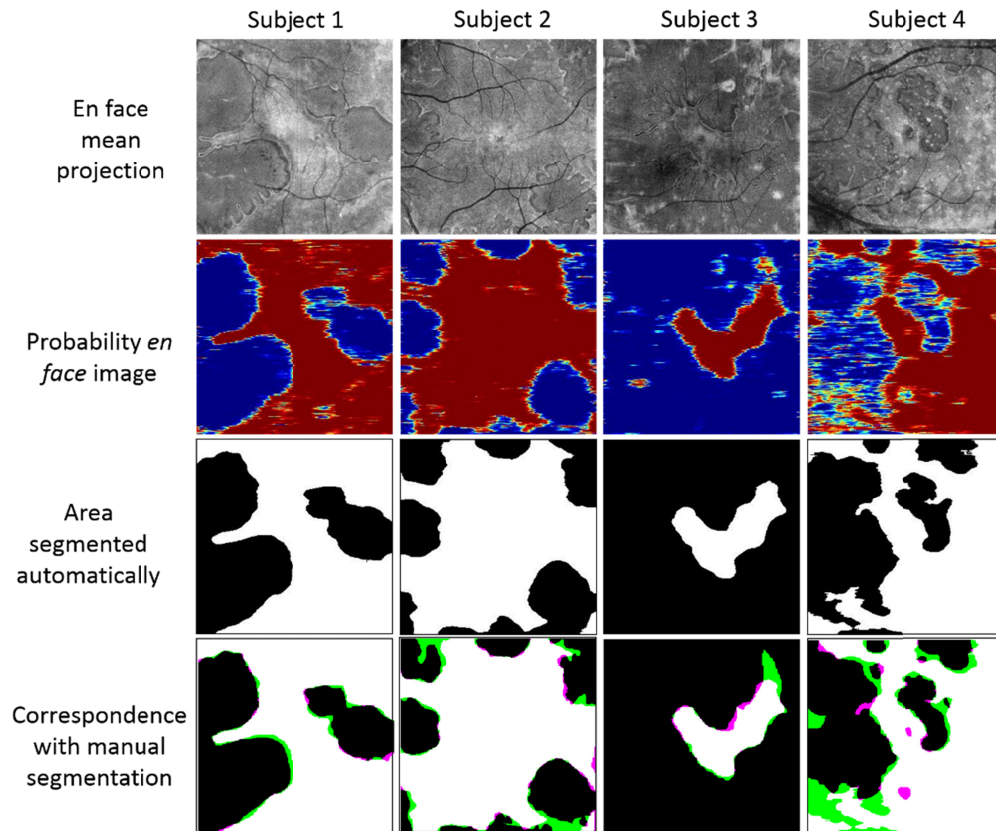


Fig. 8. Preserved ellipsoid zone areas detected automatically by the deep learning method and their correspondence with manual segmentation for 4 subjects with choroideremia. Areas in white show agreement between manual segmentation and the CNN classifier, whereas areas in green are false negatives and areas in purple are false positives.

The effect of removing the step consisting of rectification of the EZ data underneath large vessels was evaluated in five healthy subjects [Fig. 9]. It is observed that artefactual shadows underlying large vessels would be mistakenly missed from the detected areas of preserved EZ [Fig. 9(B-C)] if the vessel shadows were not removed in the pre-processing step.

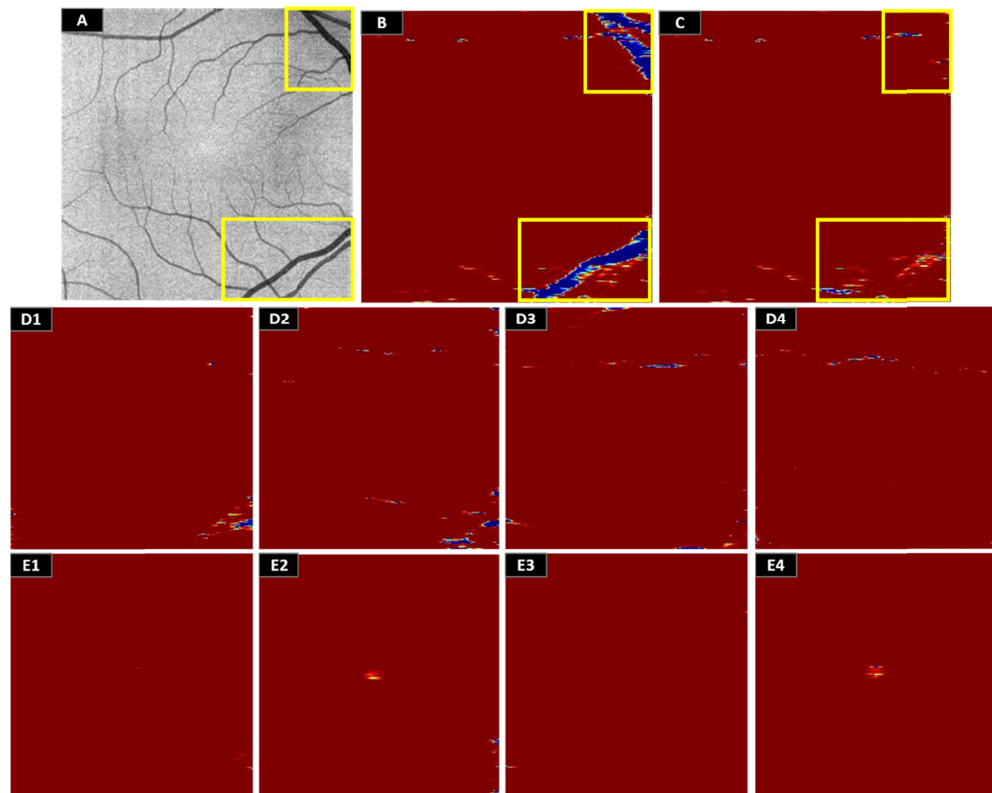


Fig. 9. Effect of large vessel shadow artifacts on the accuracy of CNN patch-wise classifier in normal data. (A) shows the en-face projection of the rectangular section selected above the Bruch's membrane. Shadow artifacts caused by absorption of superficial large vessels artifacts are evident. The CNN probabilistic images generated without rectifying shadow artifacts (B) and rectifying shadows (C) were generated and yellow boxes mark the regions where classification accuracy was improved. (D1-D4) show the classification results for the remaining four healthy subjects training with the retinitis pigmentosa data set and (E1-E4) show the classification results for the same subjects training with the choroideremia data set.

#### 4. Discussion and conclusion

We have proposed an algorithm based on neural network classification of B-scan patches for the detection of EZ defects in choroideremia and retinitis pigmentosa. The same neural network was trained and tested on data from two different IRDs, showing in both cases good agreement with ground truth. The manual segmentation method used for annotations was different for each IRD (EZ thickness map for retinitis pigmentosa and mean EZ slab projection for choroideremia), exploiting the *en face* characteristics of each disease that provide best contrast to a human grader. After a 2D binary mask of the preserved EZ area was generated, A-lines were labeled according to their (x, y) position.  $33 \times 33$ -pixel patches were generated from B-scans sections right above the Bruch's membrane interface, and were categorized as preserved/loss EZ according to the classification of its central A-line. Then, the trained CNN could generate two-dimensional *en face* probability maps of the EZ loss/preserved area. Whereas data acquired by the Angiovue OCTA system was used for software development, this method is applicable to any OCT scanner commercially available.

In this paper we have used 10 scans of diseased eyes for network training and tested on a different set of 10 scans. In order to provide a large training data set to the deep learning model, we divided B-scans into small patches. Although the training data provided a total of 57120 patches and the algorithm's performance on the test sets was satisfactory, the training

data did not contain enough representation of inter-subject variability and pathological appearances. Even, a significant difference from ground truth was observed in one case with retinitis pigmentosa [Fig. 10] in the region with partial EZ preservation. The small set of diseased subjects was unlikely to contain all features by which the disease manifests in OCT B-scans of the whole population of diseased subjects, resulting in inadequate learning of this particular feature. For solutions targeting diseases with such a small prevalence in the population, accessibility to larger databases product of inter-institutional collaboration would be beneficial in order to exploit the fullest potential of deep learning.

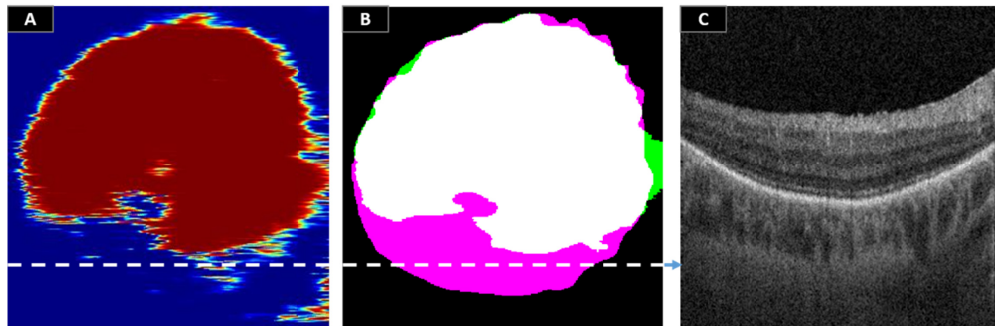


Fig. 10. Case with the greatest difference in ellipsoid zone area between the CNN algorithm and manual grading. A is the probability image of the CNN output; B is the comparison of automatic and manual results, the areas in green are false negatives and areas in purple are false positives; C is the B-scan image in the dashed line position.

The preserved EZ zone in the disease of choroideremia has the peculiarity of presenting a completely preserved area (very hyper-reflective in *en face* projections such as in Fig. 8) and a partially preserved area with lower reflectivity surrounding it. Since the partially preserved area is hard to segment *en face*, in a previous work where we trained a random forest for *en face* EZ detection in choroideremia alone, it was necessary to generate a total of 12 feature maps to recognize preserved from EZ loss with accuracy. The deep learning approach proposed here simplifies significantly our previous machine learning solution, as it is trained with sections of the B-scans themselves rather than *en face* projections, and no subjective selection of features is required. Although this characteristic makes the deep learning approach more robust, neither random forests nor deep learning could detect the outer retinal tubulations protruding from the main preserved EZ area. These tubulations with pseudopodial appearance are typical of choroideremia and have been attributed to an outer-retina scrolling surviving mechanism after losing trophic support from underlying RPE and choriocapillaris [29]. When the network was trained asking the grader to include tubulations in the manual segmentation the network performance was worse by 10%. We attribute the decreased performance to three factors. First, the loss of RPE under the partially hyper-reflective EZ at tubulation positions. Second, the fact that these tubulations are very thin and in many patches manually classified as preserved EZ the majority of the A-line positions would in fact be EZ-loss. Third, the inability of graders to account for all tubulations (Fig. 11), hence feeding the network with some tubulation patches classified as preserved EZ and some others as EZ loss. The spatial overlap with choroidal flow loss observed in our previous investigation suggests that the positions of these outer retinal tubulations are likely among the photoreceptor areas to be lost next in the natural progression of the choroideremia disease [18]. Currently, we add the image post-processing technique based on a local active contour routine proposed previously [18] in order to detect most of the pseudopodial extensions.

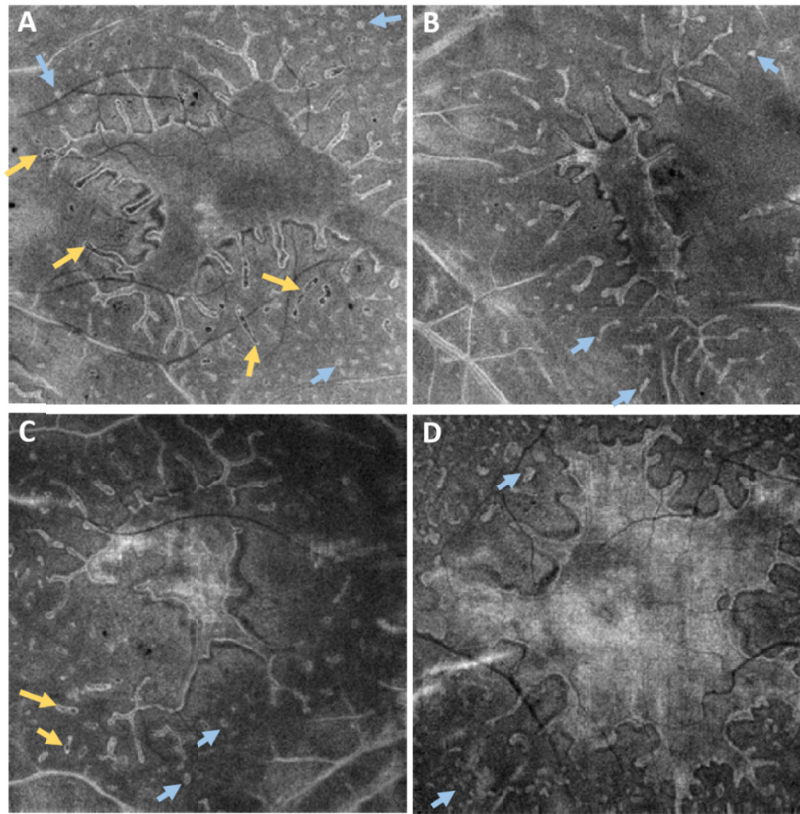


Fig. 11. Examples of *en face* projections of the choroideremia cases that illustrate the challenges of manually segmenting the outer retinal tubulations. Isolated areas possibly caused by vanishing tubulations (blue arrows) are numerous, of small size and scattered over the whole scanning area. They are difficult to segment manually and likely responsible of misclassifications when included in the gold standard. Some tubulations have dark appearance within the boundaries apparent to a human grader (yellow arrows). Such patches exhibit an appearance very different to the typical preserved EZ patch observed in Fig. 5.

A similar deep learning model has been used previously in Ref [30], for the task of retinal layer and drusen segmentation in age-related macular degeneration. Unlike Fang et al., our proposal is to generate the gold standard by manual grading *en face* images (projection of a carefully selected slab in choroideremia and a thickness map in retinitis pigmentosa) rather than B-scans, demanding minimal layer segmentation requirements. Accurate manual classification of the input data is critical in the proper training of supervised machine learning methods. In diseases where EZ loss can be partial and the quality of B-scans is often low, there is more confidence for human graders to draw the boundaries of the damaged/preserved area from *en face* images (Fig. 12). Moreover, there is no context available from neighboring regions to a B-scan manual grader, potentially making that method more prone to confusion at boundaries. By using an OCT system with a significantly denser B-scan sampling compared to the system used by Fang et al, we could confidently use the labels generated *en face* to classify underlying patches of cross-sectional B-scans.



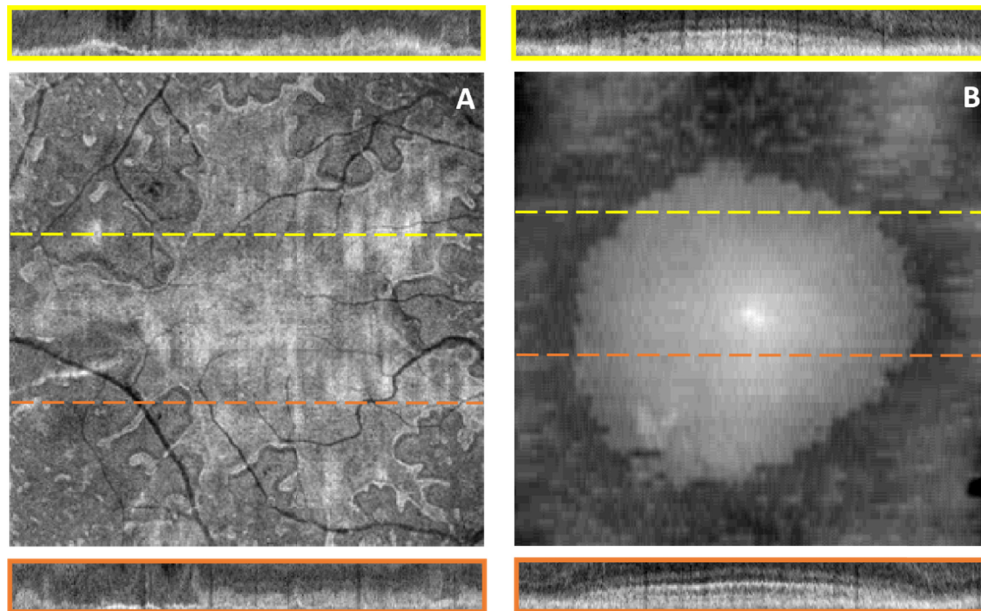


Fig. 12. Comparison of the contrast between preserved EZ and EZ loss in *en face* vs cross-sectional images in choroideremia (A) and retinitis pigmentosa (B). Dashed lines indicate the positions at which 33-pixel-width sections of B-scans above the Bruch's membrane are extracted, which are represented enclosed in yellow and orange boxes. Manual segmentation of the boundary between the two classes can be performed with better confidence from the *en face* projection (choroideremia) or thickness map (retinitis pigmentosa) compared to the B-scan sections.

In summary, we have used a single deep learning platform to automatically detect EZ loss in choroideremia and retinitis pigmentosa. Patch-wise training the CNN for classification of segments of A-lines (represented by a pixel in *en face* images) solved the preserved EZ segmentation problem. Although this work has only been performed on two IRDs thus far, it has potential to be trained to manage many IRDs simultaneously owing to the flexibility of the deep learning method.

### Funding

National Institutes of Health (Bethesda, MD) (R01EY027833, DP3 DK104397, R01 EY024544, P30 EY010572); National Natural Science Foundation of China (NO. 61471226); Natural Science Foundation for Distinguished Young Scholars of Shandong Province (NO. JQ201516); China Scholarship Council, China (grant no.: 201608370080); unrestricted departmental funding grant and William & Mary Greve Special Scholar Award from Research to Prevent Blindness (New York, NY).

### Acknowledgments

The authors also thank the support from Taishan scholar project of Shandong Province.

### Disclosures

David Huang: Optovue, Inc (F, I, P, R). Yali Jia: Optovue, Inc (F, P). These potential conflicts of interest have been reviewed and managed by OHSU. Other authors declare that there are no conflicts of interest related to this article.